

平成 30 年度厚生労働科学研究費補助金  
(政策科学総合研究事業(臨床研究等 ICT 基盤構築・人工知能実装研究事業))  
「介護施設入居高齢者等の疾病の早期発見・重症化予防をAIを活用して行う実証研究」  
総括研究報告書

研究代表者:

今中雄一 (京都大学大学院医学研究科医療経済学分野 教授)

研究分担者:

鹿島 久嗣 (京都大学大学院情報学研究科知能情報学専攻 教授)

櫻井 保志 (大阪大学産業科学研究所トランスレーショナルデータビリティ研究分野 教授)

國澤 進 (京都大学大学院医学研究科医療経済学分野 講師)

研究協力者:

佐々木典子 (京都大学大学院医学研究科医療経済学分野 特定准教授)

林 慧茹 (京都大学大学院医学研究科医療経済学分野 研究員)

原 広司 (京都大学産官学連携本部 特定助教)

中部 貴央 (京都大学大学院医学研究科医療経済学分野)

寺岡 英美 (京都大学大学院医学研究科医療経済学分野)

要旨

目的:

介護施設等に居住する高齢者等の疾病の早期発見・重症化予防を行うために、各種データを用いた評価・通知のシステムを研究開発し、現場にフィードバックすることを目的としている。

方法:

- 1 【生体センサーデータ】夜間の見守りセンサーデータを用いて、入居者の状態をモデル化する。
- 2 【介護提供組織の体制・風土データ】介護の質との関連が深い介護提供者の組織文化を明らかにし、組織レベルでの悪化リスクを定量化する。
- 3 【健康関連データ(介護レセプト、調査票データ、介護カルテ等)】
  - 3.1 入居者の QOL:QOL をモニター・把握し、QOL 変化を予測する。
  - 3.2 レセプト等の情報の活用:地域在住高齢者認知症リスクスコア評価尺度の作成、介護サービス利用者のアウトカム予測及び自己負担変化から与えた影響を検証する。
  - 3.3 介護カルテ情報の活用:介護カルテから入居者の状態変化をモニター・把握し、状態悪化を予測する。

結果:

1 【生体センサーデータ】

使用したデータは、ある介護施設の A から L までの 12 人の入居者から 3 日間にわたって計測された、心拍数、呼吸数、呼吸レベルの 3 次元のセンサで構成されている。提案手法は複数入居者のセンサデータから、退室や歩行、睡眠などの入居者ごとに異なる活動状態を抽出し、共通の状態ごとにモデル化することに成功した。

## 2 【介護提供組織の体制・風土データ】

介護ケアのパフォーマンスは組織により大きくばらつきと想定され、人が直接ふれあう介護のパフォーマンスにおいては、特に人と組織の影響が大きいと考えられる。組織の力は、要介護者の状態悪化予防に大きく寄与する可能性があり、組織風土の重要性は高い。

かつて開発した病院職員を対象とした組織文化調査質問紙を介護版に改訂した。改訂版調査票の統計的な信頼性・妥当性を検証するため、クロンバック $\alpha$ および確認的因子分析を用いた。その結果、改訂版調査票の信頼性・妥当性が確認され、その成果を学会で発表した。

本調査票での組織文化は、「チームワーク」や「情報共有」、「士気・やる気」、「プロとしての成長」、「組織の価値観」、「ケアの質を支える資源」、「責任と権限」、「利用者安全の改善のシステム」の8つの領域から構成されている。領域ごとに質問項目を4~7問用意し、それらを領域別にスコアリングしている。さらに、職場環境に関わる領域として「職務満足度」や「仕事量と負担」なども設問に加えた。既存研究において、組織文化スコアが高い病院では、医療の質指標も高いことが明らかになっている。

2018年4月からのおよそ1年間で、7法人、77の介護事業所、1355人を対象に調査を実施し、1069人から回答を得た。同一法人内であっても、施設によって組織文化スコアに大きなばらつきがみられた。つまり、法人レベルだけでなく、施設レベルでも介護の質にばらつきがみられる可能性が示唆された。職員のヒアリングにより、施設によって研修の頻度や内容、運用方針等に違いがあり、そういった要因が関連している可能性がある。

職位間で比較すると、中間管理職が、幹部や非管理職に比べて組織文化スコアが低い傾向がみられた。とくに「資源(の充足感)」や「(組織としての)改善のシステム」、「職務満足度」等で中間管理職のスコアが低い傾向にあった。介護施設でのヒアリングを通じて、中間管理職の確保および育成に課題があることが指摘されており、今回の結果はそれを支持するものであった。介護施設における中間管理職の確保・育成の取組がうまく機能している組織では、介護の質も高い可能性が示唆された。

また、介護分野では人手不足が極めて大きな問題になっており、職員の確保、リテンションマネジメントを検討することが重要である。本調査のデータにおいて、職員の職務満足度や職場への定着意欲と関連が深い領域を、重回帰分析を用いて検証した。その結果、「プロとしての成長」や「責任と権限」、「仕事量と負担」、「(組織の)将来像」がこれらとの関連していることが明らかになった。とくに、「プロとしての成長」の偏回帰係数が最も高く、関連が強いことが示された。研修機会の提供や、日々の業務の中での技術的な指導、職員間で学び合う環境づくりを作り出すことで、職員のリテンションマネジメントにつながると考えられる。この成果の一部を学会で発表した。

## 3 【健康関連データ(介護レセプト、調査票データ、介護カルテ等)】

### 3.1 入居者のQOL

入居者のQOLは、生活をする場である介護施設において、重要なアウトカムの一つである。当研究チームでは、入居者のQOLをモニター・把握し、その変化を予測することを目指している。

入居者のQOL等を把握するため、入居者の「生活とケアの満足度」調査票を開発した。調査項目は、世界で最も使用されているQOL尺度のEQ-5D、精神的健康状態を把握するために

WHO-5、幸福感、介護サービスに対する満足度等とした。

2018年4月から調査を開始し、上記の組織文化調査と同時に調査を行った。5法人、約50事業所、2677人を対象とし、そのうち1701人から回答を得た。EQ-5DやWHO-5といった世界的に使用されている指標について、介護サービス利用者の状況はこれまでほとんど明らかになっておらず、本調査によってこれらの基礎的なデータを取得できた。EQ-5Dでは、介護サービスを利用していない一般の70代以上の人で0.866だが、利用者では0.48～0.56であることが明らかになった。また、施設サービス利用者のほうが在宅での介護サービス利用者よりもEQ-5Dは高い傾向がみられた。ただし、対象法人が限られているため、さらなる検証が必要である。同様に、WHO-5では、一般の70代前半は16.9に対して、介護施設利用者は15.1、在宅での介護サービス利用者は11.6であった。その他の年代でも一般に比べて介護サービス利用者のほうが低く、かつ在宅での介護サービス利用者のほうがさらに低い傾向にあった。こうした基礎的なデータは、利用者の実態を把握するうえで重要な情報であり、今後はこの変化を把握し、その予測モデルを構築できるように、引き続き調査や分析を進める。

職員の組織文化調査の結果と、利用者の生活とケアの満足度調査の結果の相関関係を検証した。その結果、利用者のサービス満足度と職員の「チームワーク」との間の関連が強いことが明らかになった。ただし、事業所単位での分析となるため、検証にはサンプル数をさらに増やす必要があり、また施設ごとの利用者の状態に違いがあることがこれらの結果に影響している可能性がある。今後、個人や施設ごとに調整をし、利用者の重症化予測につなげる。

### 3.2 レセプト等の情報の活用

重症化予測の知見を生み出すため、AI・機械学習および大規模データと調査票を活用し、以下の分析を行った。

第1に、要介護認定時の基本チェックリストおよび健診データを用いて、Cox比例ハザードモデルの結果から認知症リスクスコア評価尺度を作成した。第2に、介護サービスレセプトデータを用いて、「差分の差分法」(difference-in-difference estimation)によってその介護自己負担の変化が居宅サービス利用時間と介護費用それぞれに与える影響を検証した。第3に、AI・機械学習を活用した予測モデルによって、介護サービス利用高齢者の死亡予測モデルを作成した。一都道府県で、介護レセプトデータベース、国民健康保険データベース、後期高齢者医療保険データベースを連結、追跡した分析を行った。観察期間中に新規診断された疾病について、RandomForestを用いて各要因と死亡との関連重要度 Mean Decrease Accuracy を明らかにして予測モデルに適用した。その結果、介護状況と疾病から、極めて高い予測力を実現した。第4に、要介護度重症化予測ではDeep Learningを用いて、要介護度別の介護サービス利用の組み合わせによる、利用者の要介護度変化に与える影響を解明することを目的とした分析を行った。一都道府県の介護レセプトデータベースで、要介護度別のサービス利用をDyadic Soft Clustering分析を行った。Dyadic Soft Clusteringで介護サービス利用パターンを10グループを時系列で示した。

### 3.3 介護カルテ情報の活用

介護カルテの事業を行っている株式会社介護サプリアから介護記録データを研究のために提供を受けた。このデータには、言語記録データが約 470 万件、200 事業所、約 3000 人分、2 年間分のデータが含まれている。データ項目には、利用者基本情報(性別、年齢、要介護度など)、利用者バイタル情報(体温、脈拍、血圧など)、利用者飲水情報、利用者食事情報、入浴情報(入浴方法など)、排泄情報(排泄場所、量など)、服薬情報がある。

約 470 万件の自由記述データのうち、約 77 万件のデータを抽出し、「熱」および「転倒」を含むデータにフラグを立てた。「熱」というキーワードを含むデータは 7572 件(出現率 0.98%)だった。実際に発熱が確認されたのは 8646 件(出現率 1.12%)だった。「転倒」を含むデータは 2193 件(出現率 0.28%)であり、実際に「転倒」が確認されたのは 831 件(0.11%)だった。

「熱」および「転倒」に関するソースを整理し、およその出現率を確認した。今後、このデータを機械学習等に活用し、解析を進める。

結語:

当研究は、現時点でデータ収集と解析試行成果が基盤として整ってきており、一部成果を発表した。最終年度の次年度では、これらの成果をまとめていく。具体的には、以下の点を進める。

- (1) 生体センサーデータ: 身体的活動・睡眠のパターン抽出と変化の予測、AI技術の適用
- (2) 組織体制・文化データ: 組織文化と利用者 QOL との関連性の実証
- (3) 入居者健康関連データ(介護レセプト、介護カルテ、調査票等): 状態把握と悪化リスク要因の同定
- (4) 統合した悪化予測モデルの構築とその現場利用モデルの概念実証

## A. 目的

AIを用いることで、介護施設等に居住する高齢者等の疾病の早期発見・重症化予防を行うために、各種データを用いた評価・通知のシステムを研究開発し、現場にフィードバックすることを目的としている。

### 1)【生体センサーデータの解析】

近年の IoT デバイスの急速な普及に伴い、それらのデバイスから収集した多様かつ大量のデータを管理、解析することにより、高度なサービスに活用しようとする動きが盛んである。医療介護分野においては、ビッグデータ解析は医療介護サービスの質の向上および効率化を図り、様々な問題を解決できる重要なアプローチとして期待されている。本研究の目的は、介護施設入居者から得られた多種多様なセンサデータから、入居者の状態をモデル化し、入居者の状態や特徴を抽出、分類、さらに

予測を行うことを可能とするデータ解析のための AI 関連技術を開発することである。

### 2)【介護提供組織の体制・風土データ】

介護提供者の組織風土が、介護施設入居高齢者の健康状態・活動状態に大きく影響すると考えられる。医療においては、組織風土に関する調査がいくつか行われているが、介護ではそれらがほとんど明らかになっていない。そこで、本研究では、介護施設のサービス向上・改善に活用することを目指し、介護提供者の組織風土を計測することで、組織基礎リスクの定量化することを目的とする。

### 3)【健康関連データ(介護レセプト、調査票データ、介護カルテ等)】

#### 3.1)入居者の QOL

入居者の QOL や精神的健康状態、幸福感、サービス満足度等を測定し、それらをモニター・把握することで、サービスの向上や変化の予測を目的と

する。

### 3.2)レセプト等の情報の活用

認知症リスクスコア評価尺度の作成、及び AI・機械学習を応用して、介護・医療レセプト等の大規模データや調査票データを活用した要介護度重症化・死亡率予測モデルの構築、介護自己負担の変化から介護利用者に与えた影響を検証する。

さらに、データマイニングの技術・手法などで、年齢、性別、地域別、傷病、入院日数、検査、薬剤、処置、介護サービス、要介護度などを変数として用い、複数期間において、年齢、性別、傷病、利用した介護サービス項目、要介護度などを説明変数とし、パネルデータ分析を行い、リスクが高い群を同定する。また、Deep Learning を利用して、要介護度別に利用したサービスを Dyadic Soft Clustering した結果と、性別、年齢、保険者を説明変数とし、一年後の要介護度を予測するモデルを構築する。上述のモデルの説明変数と一年後要介護度重症化有無をアウトカムにした予測モデルを構築、さらに従来の回帰モデルを同じデータに用い、モデルの精度を比較した。

### 3.3)介護カルテ情報の活用

介護記録には、利用者の状態変化(バイタルや転倒・発熱などのイベントの発生等)について記録されている。このデータを活用し、利用者の重症化の予測モデルを構築することを目的とする。データの整形およびテキスト抽出のためのシソーラスの構築を目指す。

## B. 対象・方法

### 1)【生体センサーデータの解析】

本研究で扱うデータは、{入居者, センサ, 時間}の複数ドメインを持つテンソルとなっており、入居者の病状や特徴を多角的に解析する必要がある。具体的には、(1) 食事、読書、就寝など入居者の

行動を表す時系列パターンとその変遷、(2) 入居者毎の共通パターン/入居者個人特有のパターンの抽出である。特に後者は(2)は、入居者間の行動の違いのみならず、就寝のような同じ行動であっても異なる振る舞いを示すような、入居者毎の特徴を示すものである。(1)の状態の変遷から病状悪化の予兆を検知したり、(2)により異常な振る舞いを示す入居者を早期に発見し、より細かく状態を観察することが可能になる。

以上を踏まえ、本研究目的を達成するため、大規模介護データのための多角的解析技術の開発を行い、入居者の状態を多方面から分析する。

**提案手法：**提案手法の概要を図1に示す。提案手法は時間方向の解析を行う V-Split と、入居者方向の解析を行う H-Split で構成される。V-Split では、入居者の時間方向の状態遷移をセグメント分割し、共通セグメント(ここではレジームという)をモデル化する。一方 H-Split は、V-Split で得られたレジームの中から、入居者ごとの違いを抽出し、別のレジームとして表現する。これら2つのアルゴリズムを任意の順序で繰り返し、最終的な解を求める。

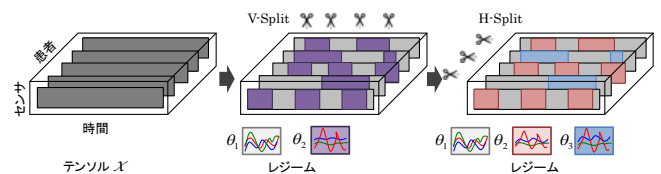


図1. 提案手法の概要

### 2)【介護提供組織の体制・風土データ】

介護事業所の職員に対して、自記式質問紙を配布し、無記名で調査票を封筒に入れ、厳封したうえで事業所ごとに回収をした。調査対象者は、介護事業所に勤務するすべての者(介護士、看護師、調理、清掃等)とした。質問紙には、回答者の職種や職位などを記入する欄を用意し、職種・職位ごとの違いも検証できるようにした。

質問項目の妥当性を確認するため、専門家を交えて内容を検証した。次に、作成した調査票の統

計的な信頼性・妥当性を検証するため、2 法人約 200 人に対してパイロット調査を実施した。結果を踏まえて、質問項目を一部改訂した。

統計的な妥当性・信頼性が確認された質問票を用いて、先の 2 法人に加えて、新たに 5 法人で調査を実施した。調査対象者は合計で約 1300 人となった。回収したデータを用いて、記述統計や職務満足度・勤務継続意欲を目的変数にした重回帰分析を行った。

### 3)【健康関連データ(介護レセプト、調査票データ、介護カルテ等)】

#### 3.1)入居者の QOL

協力を得られる介護事業所の利用者に対して、自記式質問紙を配布し、無記名で調査票を封筒に入れ、厳封したうえでポストに投函してもらい、調査票を回収した。ただし、重度な認知症患者等の回答が極めて困難な利用者については、施設側の判断で除外した。利用者が回答の記入を困難としている場合は、家族が補助をして回答をしていただいた。家族の訪問が少ない利用者の場合は、職員が回答の補助を行った。

調査票には、QOL、精神的健康状態、幸福感、サービス満足度等の項目を含んでいる。QOL は EQ-5D(世界的に使用されている QOL 指標の一つ)、精神的健康状態は WHO-5(WHO が開発した精神的健康指標)を用いた。幸福感は、幸福感研究の第一人者である京都大学こころの未来研究センターの内田由紀子氏に協力いただき、項目を設定した。サービス満足度は、先行研究および専門家を交えて設定し、実際の介護職員へのヒアリングおよび利用者へのヒアリングにより改訂を行った。

5 法人約 55 事業所に協力いただき、対象利用者数はおよそ 2600 人だった。

#### 3.2)レセプト等の情報の活用

A 市に在住する 65 歳以上の高齢者において基本チェックリストや健診データを用い、2011 年から

2015 年までの約 4 年間の追跡により、新たに要支援・要介護認定された方の中で認知症が発症しているかどうかを分析し、認知症リスクスコア評価尺度を作成した。

A 県より 2014 年 8 月から 2016 年 7 月に 65 歳以上かつ要介護度 1 以上の合計 954,047 人・月介護サービスレセプトデータを用いて、性、年齢、要介護度、補助の有無などを共変量として、「差分の差分法」(difference-in-difference estimation)により、その介護自己負担の変化が居宅サービス利用時間、施設入居日数と介護・医療費用それぞれに与える影響を検証した。

B 県の介護保険、国民健康保険、後期高齢者医療制度のレセプトデータを連結し用いた。2010 年 10 月から 2011 年 9 月の間に、介護サービス利用があり、かつ 2011 年 10 月に要介護度が明らかであり、さらに 2014 年 9 月まで追跡できた者あるいは死亡した者を対象とした。2010 年 10 月を登録月とし、2009 年 10 月から 2010 年 9 月までに診断された疾患を ICD10 コードより同定し、それを既存疾患という説明変数とし、登録月に新たに診断された疾患を新規疾患、登録月の介護サービス利用種類、登録月の入院日数を説明変数とした。ランダムフォレストで変数の既存疾患と新規診断された疾患の重要度を示して、そして高齢者死因ランキング上位の疾患及び介護が必要になった原因となる主な疾患を考慮した上で説明変数を選択した。

年齢、性別、独居および上記で選択された説明変数を用いて、Kaplan-Meier 生存分析を行った。登録日から観察終了月(2014 年 9 月)までの約 3 年間の死亡有無を目的変数とし、Cox 回帰分析で死亡予測モデルを構築した。

要介護度予測及び重症化予測モデルは、N 県で、2017 年 4 月から 2018 年 3 月までの介護レセプトデータベースで、要介護度別のサービス利用を Dyadic Soft Clustering 分析を行った。結果と最も関連がみられるグループと性別、年齢、保険者を Deep Learning に適用して、2018 年 3 月の要介護



度および重症化有無を予測した。さらに、同じデータを用い、従来の回帰分析モデルを作成して、モデルの精度を比較した。

### 3.3)介護カルテ情報の活用

株式会社介護サプリアおよび介護サプリアの電子カルテを利用する介護事業所のデータ提供を受けた。介護サービス形態は様々であり、有料老人ホームやサービス付き高齢者向け住宅などである。対象者数は約 2000 人、約 2 年間のデータが含まれている。

データには、データ項目および言語情報があり、具体的には、個人属性、日時情報、バイタル、飲食・排泄、異変の把握がある。異変の把握とは、介護者(介護士や看護師等)が記録した自由記述データを指す。およそ 470 万件のデータが含まれている。この記録には、発熱や転倒、入院、不穏等のイベントが記載されるため、それらのキーワードを拾い上げて、アウトカムを把握する。ここで把握されたアウトカムを目的変数とし、個人属性で調整をして、バイタル、飲食・排泄、その他の自由記述の内容を説明変数とした予測モデルの構築を目指す。そのために、まずは異変の把握からアウトカムを適切に抽出するためのシソーラスを構築する。

## C. 結果

### 1)【生体センサーデータの解析】

実データに対する実験結果を図 2 に示す。

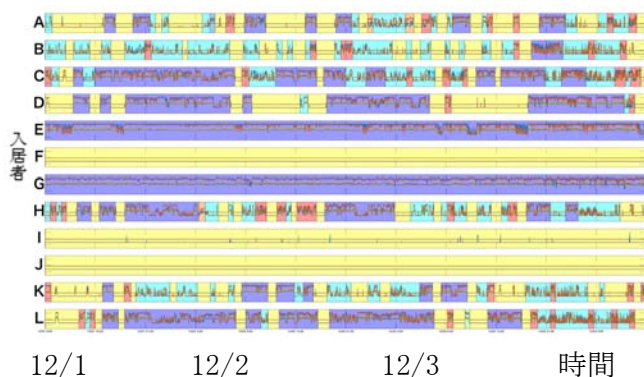


図 2. 実データに対する実験結果

使用したデータは、ある介護施設の A から L までの 12 人の入居者から 3 日間にわたって計測された、心拍数、呼吸数、呼吸レベルの 3 次元のセンサーで構成されている。提案手法は複数入居者のセンサーデータから、退室や歩行、睡眠などの入居者ごとに異なる活動状態を抽出し、共通の状態(図 2 における同じ色のセグメント、すなわちレジーム)ごとにモデル化することに成功した。

### 2)【介護提供組織の体制・風土データ】

2018 年 4 月からのおよそ 1 年間で、7 法人、77 の介護事業所、1355 人を対象に調査を実施し、1069 人から回答を得た。同一法人内であっても、施設によって組織文化スコアに大きなばらつきがみられた。つまり、法人レベルだけでなく、施設レベルでも介護の質にばらつきがみられる可能性が示唆された。職員のヒアリングにより、施設によって研修の頻度や内容、運用方針等に違いがあり、そういった要因が関連している可能性がある。

職位間で比較すると、中間管理職が、幹部や非管理職に比べて組織文化スコアが低い傾向がみられた。とくに「資源(の充足感)」や「(組織としての)改善のシステム」、「職務満足度」等で中間管理職のスコアが低い傾向にあった。介護施設でのヒアリングを通じて、中間管理職の確保および育成に課題があることが指摘されており、今回の結果はそれを支持するものであった。介護施設における中間管理職の確保・育成の取組がうまく機能している組織では、介護の質も高い可能性が示唆された。

また、介護分野では人手不足が極めて大きな問題になっており、職員の確保、リテンションマネジメントを検討することが重要である。本調査のデータにおいて、職員の職務満足度や職場への定着意欲と関連が深い領域を、重回帰分析を用いて検証した。その結果、「プロとしての成長」や「責任と権限」、「仕事量と負担」、「(組織の)将来像」がこれらとの関連していることが明らかになった。とくに、「プロとしての成長」の偏回帰係数が最も高く、関連が

強いことが示された。研修機会の提供や、日々の業務の中での技術的な指導、職員間で学び合う環境づくりを作り出すことで、職員のリテンションマネジメントにつながると考えられる。この成果の一部を学会で発表した。

### 3)【健康関連データ(介護レセプト、調査票データ、介護カルテ等)】

#### 3.1)入居者の QOL

入居者の QOL は、生活をする場である介護施設において、重要なアウトカムの一つである。当研究チームでは、入居者の QOL をモニター・把握し、その変化を予測することを目指している。

入居者の QOL 等を把握するため、入居者の「生活とケアの満足度」調査票を開発した。調査項目は、世界で最も使用されている QOL 尺度の EQ-5D、精神的健康状態を把握するために WHO-5、幸福感、介護サービスに対する満足度等とした。

2018 年 4 月から調査を開始し、上記の組織文化調査と同時に調査を行った。5 法人、約 50 事業所、2677 人を対象とし、そのうち 1701 人から回答を得た。EQ-5D や WHO-5 といった世界的に使用されている指標について、介護サービス利用者の状況はこれまでほとんど明らかになっておらず、本調査によってこれらの基礎的なデータを取得できた。EQ-5D では、介護サービスを利用していない一般の 70 代以上の人で 0.866 だが、利用者では 0.48～0.56 であることが明らかになった。また、施設サービス利用者のほうが在宅での介護サービス利用者よりも EQ-5D は高い傾向がみられた。ただし、対象法人が限られているため、さらなる検証が必要である。同様に、WHO-5 では、一般の 70 代前半は 16.9 に対して、介護施設利用者は 15.1、在宅での介護サービス利用者は 11.6 であった。その他の年代でも一般に比べて介護サービス利用者のほうが低く、かつ在宅での介護サービス利用者のほうがさらに低い傾向にあった。こうした基礎的なデータは、利用者の実態を把握するうえで重要な情報であり、

今後はこの変化を把握し、その予測モデルを構築できるように、引き続き調査や分析を進める。

職員の組織文化調査の結果と、利用者の生活とケアの満足度調査の結果の相関関係を検証した。その結果、利用者のサービス満足度と職員の「チームワーク」との間の関連が強いことが明らかになった。ただし、事業所単位での分析となるため、検証にはサンプル数をさらに増やす必要があり、また施設ごとの利用者の状態に違いがあることがこれらの結果に影響している可能性がある。今後、個人や施設ごとに調整をし、利用者の重症化予測につなげる。

#### 3.2)レセプト等の情報の活用

約 4 年間のデータで 72,127 名地域在住高齢者を追跡し、そのうち 6,656 人(約 9.2%)が新たに要介護・要支援の認定され、認知症を発症した。基本チェックリストの中の 13 項目と、性別、年齢と、健診の BMI と血糖値から、0～62 点(点数が高いほど高リスク)となる評価尺度を作成した。

2015 年 8 月から、約 1 割の介護サービス利用者の自己負担割合が 1 割から 2 割に増加した。所得による介護自己負担 1 割から 2 割に増え介護利用者は、自己負担上昇前後一年の月平均居宅サービス利用時間、施設入居日数と介護費用の変化を比較することで、相対的な月平均居宅サービス利用時間の減少が顕著に観察された。また、月平均介護費用と介護点数は大幅に減少したことが分かった。

介護利用者の死亡予測について、対象者の平均追跡期間は 959 日であり、新たに診断された疾患のない介護利用者の平均追跡期間は 962 日であったが、新たに診断された疾患があった介護利用者の平均追跡期間は 696 日だった。

男性、高齢、高い要介護度、登録月に 3 週間以上介護施設に入所・入居、および登録月の前年に入院日数が長い期間だった高齢者は、3 年間死亡に有意に正の関連があった。



既存疾患には急性心筋梗塞と大腿骨骨折以外、すべて 3 年間死亡と関連があった。新たに診断された疾病の中で、気管支及び肺の悪性新生物が最もリスクが高かった。既存か新規診断に関わらず、がんは 3 年間死亡に最もリスク高い関連因子と示した。

Deep Learning を用い、一年後の要介護度予測及び重症化予測モデルを構築した。要介護度予測モデルの精度は 0.68、要介護度重症化予測モデルの精度は 0.79 であった。さらに、同データを従来回の回帰分析を行った結果、AUC0.612 であった。

### 3.3)介護カルテ情報の活用

約 470 万件の自由記述データのうち、約 77 万件のデータを抽出し、「熱」および「転倒」を含むデータにフラグを立てた。「熱」というキーワードを含むデータは 7572 件(出現率 0.98%)だった。ただし、「熱」は含むが、「熱中」や「熱心」、「熱唱」など、発熱とは異なるキーワードが多く混ざっており、除外すべきキーワードを整理した。また、「熱」というキーワードは含まないが、具体的な体温を記載しているデータ(たとえば 38 度)も「熱」と同程度存在した(7526 件、出現率 0.98%)。これらのデータにフラグを立て、文章を検証し、「熱が下がった」や「熱ではない」などの状態回復および否定のものを除外した結果、実際に発熱が確認されたのは 8646 件(出現率 1.12%)だった。

「転倒」も同様に、シソーラスの構築を行った。「転倒」の類語には、「転ぶ」、「倒れる」、「こける」などが含まれる。「転倒」を含むデータは 2193 件(出現率 0.28%)であり、「転ぶ(ん)」、「倒れ」、「こける(た)」を含むデータは 620 件(出現率 0.08%)だった。「熱」と同様に、状態回復や否定の文章を確認し、実際に「転倒」が確認されたのは 831 件(0.11%)だった。

「熱」および「転倒」に関するシソーラスを整理し、およその出現率を確認した。今後、このデータを機械学習等に活用し、解析を進める。

## D. 考察

### 1)【生体センサーデータの解析】

図2において、赤いレジームは特に激しい活動を表しており、入居者 C、H などで顕著にみられる。これらの入居者は赤いレジームが多いだけでなく、安静時を表す黄色いセグメントの時間も相対的に短い。このようなことから、他の入居者と比較し、活動量が多く健康的であると考えられる。実際にこれらの結果と実際の病状を照らし合わせ、入居者の健康状態をより詳しく把握することが可能であると考えられる。

現在のところ、目視によって確認できる特徴的な状態が正しく抽出できていない部分もあり、詳細な入居者の活動状態の分類には成功していない一方、大まかな活動状態までは抽出できており、さらなる技術の改善が見込まれる。

### 2)【介護提供組織の体制・風土データ】

介護ケアのパフォーマンスは組織により大きくばらつく想定され、人が直接ふれあう介護のパフォーマンスにおいては、特に人と組織の影響が大きいと考えられる。そのため、介護施設における組織風土を可視化することが重要である。本研究で実施した調査は、77 事業所、1000 人以上の介護職員に対して行っており、これほどの大規模調査はほとんど行われていない。本研究の結果から、法人間および事業所間でのばらつきがあることが認められた。こうしたばらつきが介護の質に関連していると考えられる。本調査をさらに拡大させて一般化を目指すとともに、介護の質との関連を検証することが今後の課題である。

また、介護現場における人材確保は最重要課題となっている。本研究では、職員満足度や勤務継続意欲に関連する要因として、「プロとしての成長」や「責任と権限」が重要であることを明らかにした。「プロとしての成長」は、研修機会があるかどうかや上司、同僚と高め合えるかどうかという内容を含んでいる。これは、職員不足と関連しており、職

員が少ないために研修に派遣できないという事業所も多い。つまり、「プロとしての成長」を高めるためには、研修機会の提供だけでなく、業務の効率化や組織体制の見直しといった根本的な課題解決が必要と考えられる。今後、さらに調査を展開し、好事例の情報収集も同時に行っていく。

### 3)【健康関連データ(介護レセプト、調査票データ、介護カルテ等)】

#### 3.1)入居者の QOL

QOL は医療・介護における主要なアウトカムの一つである。EQ-5D は QOL 指標として広く使用されており、様々なセッティングとの比較を可能とする。本研究は、介護サービス利用者から EQ-5D の回答を得ており、1701 名分のデータという大規模調査を実現している。EQ-5D は広く使用されているものの、高齢者、とくに要介護認定を受けた高齢者の EQ-5D は明らかになっておらず、本研究によってその基礎情報を把握することができた。

QOL は、当然ながら、介護サービスの質だけではなく、本人の心身の状態、家族・友人との交流などの影響を受ける。本研究の調査票には、QOL 以外に本人の属性、精神的健康や幸福感、他人との信頼関係等の質問項目が含まれている。現時点においては、記述統計情報のみだが、今後、介護サービスとの質との関連をさらに見出せるように、これらのデータを用いて調整をする必要がある。

#### 3.2)レセプト等の情報の活用

本研究で開発した認知症リスクスコア評価尺度は基本チェックリストと健診の項目を使っているため、非専門職でも使用でき、一般化可能性が高い。また、新規要介護認定者に対して、認知症発症に高い予測力のあるモデルが作成できた。認知症予防施策に有用なツールになりうると思われる。

また、2015年8月に介護負担が 2 割になったグループの平均要介護度と平均年齢は、1 割負担のグループよりも低かった。相対的に年齢が若い高

齢者の方が一定所得以上(収入が現役並み)、そして自立度が高い可能性がある。2 割負担グループの平均要介護度が低く、その結果は居宅サービス利用者の報告分析結果と同じ傾向がみられた。居宅サービス利用時間減少にもかかわらず、施設サービス利用は増加傾向だった。先行研究では、居宅サービスの利用により、施設サービス利用が介護費用に大きな影響を与えると指摘されている。そのため、介護費用は増加したと考えられる。

介護利用者の死亡予測において、新たに診断された疾患のうち、がんは最もリスクが高い疾患を示した。特に気管支及び肺の悪性新生物は約 3 年間の死亡リスクが大きかった。新たに疾患を診断された後の介護サービス利用者の平均追跡期間は短く、地域在住高齢者の約 3 年間の死亡に関連があった。地域在住高齢者の健康状況に影響を及ぼす新規疾病発症に対して、適切なスクリーニングが重要である。

最後に、機械学習の dyadic Soft Clustering と Deep Learning 手法を用い、一年後要介護度と重症化予測モデルを構築した。同じデータ、同じ変数で、従来の回帰モデルより高い精度が得られることが分かった。

#### 3.3)介護カルテ情報の活用

約 77 万件の自由記述データ(テキストデータ)を分析し、「熱」および「転倒」に関するシソーラスを構築した。また、それぞれの出現率も把握した。発熱はおよそ 1%、転倒はおよそ 0.1%の出現率であった。シソーラスを整理する中で、テキストデータは表記ゆれが多く、たとえば、38.5 度のピリオドが点やカンマになっているなど、様々な記載が存在する。テキストマイニングにおいては、正確かつ網羅的な辞書登録がその精度を左右するため、今回の整理を活かすことで精度を高めることができる。

今後、「熱」や「転倒」のイベント発生有無を目的変数にし、個人属性やバイタル、飲食・排泄、そのほかの自由記述から把握できるものを説明変数とし

たモデルを構築することで、介護カルテ情報を活用した重症化予測モデルの構築を目指す。

#### E. 結論

当研究は、現時点でデータ収集と解析試行成果が基盤として整ってきており、一部成果を発表した。

最終年度の次年度では、これらの成果をまとめていく。具体的には、以下の点を進める。

- (1) 生体センサーデータ： 身体的活動・睡眠のパターン抽出と変化の予測、AI技術の適用
- (2) 組織体制・文化データ： 組織文化と利用者QOLとの関連性の実証
- (3) 入居者健康関連データ(介護レセプト、介護カルテ、調査票等)： 状態把握と悪化リスク要因の同定
- (4) 統合した悪化予測モデルの構築とその現場利用モデルの概念実証

#### F. 健康危険情報

特になし

#### G. 研究発表

学会発表：

1. 中部貴央, 原広司, 今中雄一. 介護施設別・職種別の組織文化の把握および勤務継続意欲への関連要因の検討. 第 56 回日本医療・病院管理学会学術総会: 福島, 2018 年 10 月 27 日-10 月 28 日.
2. 原広司, 中部貴央, 今中雄一. 介護施設職員と病院職員の組織文化および業務・職場環境の比較—自記式質問紙を用いた横断研究—. 第 56 回日本医療・病院管理学会学術総会: 福島, 2018 年 10 月 27 日-10 月 28 日.
3. 林慧茹, 國澤進, 今中雄一. 介護保険自己負担上昇が個人レベルの総介護費用とサービス利用に及ぼす影響. 医療経済学会 第 13 回研究大会: 東京, 2018 年 9 月 1 日.
4. Honda T, Matsubara Y, Sakurai Y. Automatic Mining of Large IoT Sensor Tensor. IEEE

International Conference on Data Mining (ICDM 2018), Singapore. 17-20 November, 2018.

5. Lin HR, Imanaka Y. Predicting mortality based on the database combining health and long-term care: An indicator of quality in community integrated care systems. The International Society for Quality in Health Care 35th International Conference, Kuala Lumpur, Malaysia. 23-26 September, 2018.

論文発表：

1. Lin H, Tsuji T, Kondo K, Imanaka Y. Development of a risk score for the prediction of incident dementia in older adults using a frailty index and health checkup data: The JAGES longitudinal study. Preventive Medicine 2018;112:88-96.

図表 当研究の概要図

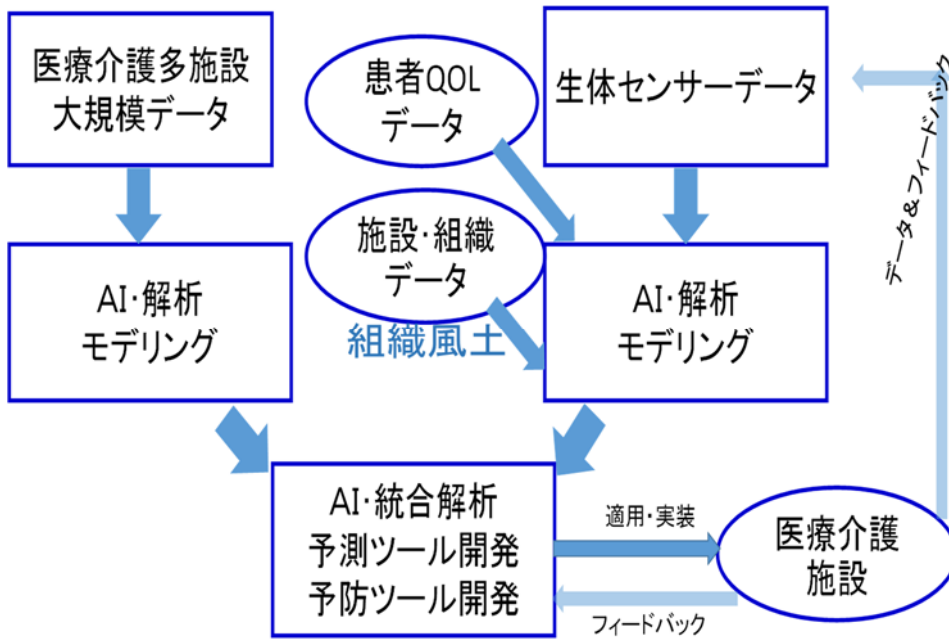


図1. 研究の流れ概要図

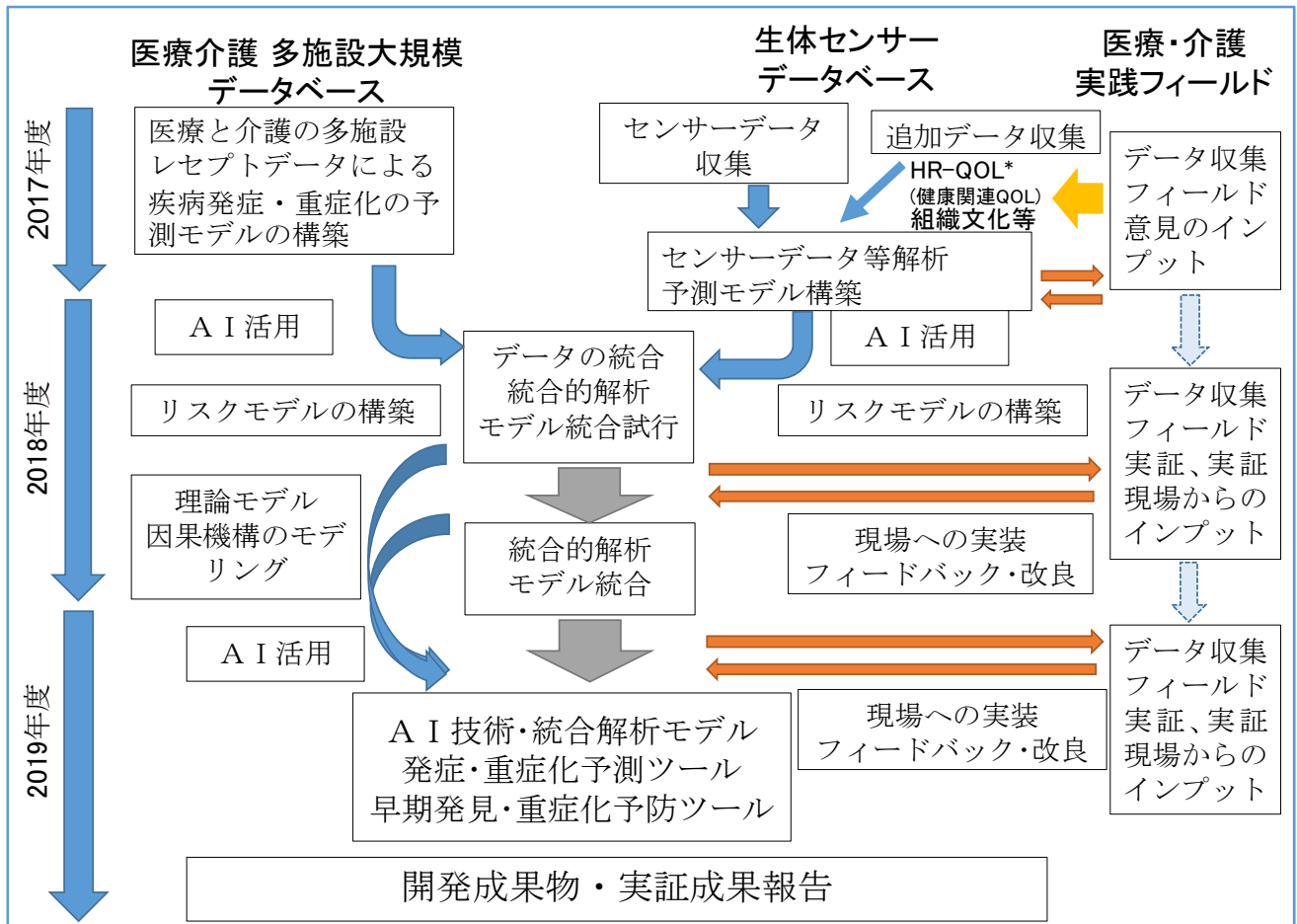


図2. 研究の流れと情報の流れ