

研究計画書

1. 研究課題

健康保険組合における健康状態および医療費の分析と疾患リスクの将来予測

2. 研究の概要

2・1 目的

西欧における先行研究では、糖尿病に起因した医療費は、非糖尿病患者のものと比較して、1.5 - 4.0 倍にもなるとの報告がある¹⁻²。主な理由としては、糖尿病に起因した合併症（糖尿病性腎症、糖尿病性網膜症、糖尿病性神経障害など）による医療費が非常に高くなることが挙げられ³、社会保障給付費が増加傾向にある⁴我が国にとっては喫緊の課題である。

本研究では、研究対象である健康保険組合集団における被保険者の健康状態を分析することで、糖尿病やその他疾患の状態を把握し、得られた分析結果をもとに糖尿病の将来疾患リスクを予測するモデルを作成することをを行う。糖尿病の疾患リスクとしては、BMIや体重増加量に代表される肥満が米国の先行研究にて指摘されている⁵。当該研究では、BMI 35 kg/m²以上で約42倍の疾患リスク（BMI 23 kg/m²未満を基準）が、体重増加 15kg 以上で約9倍の疾患リスクがある（体重増加 -2 ~ +2 kgを基準）と報告されている。また、日本人は欧米人と比較して、インスリン分泌能力が低く、軽度の肥満でも糖尿病を発症しやすいのではないかと指摘もある⁶。

これらの研究結果を踏まえた上で、本研究では、対象とする被保険者集団における医療費の要因分析および糖尿病の将来疾患リスクを分析し、新たな発見・精緻な予測モデルの構築を通し、早期介入のための判断基準となることを目的とする。また、構築した予測モデルの医療費削減効果を検証する手法を開発することで、施策の効果検証の枠組みの一助となることも目的とする。

¹ Oliva J, Lobo F, Molina B, Monereo S. Direct health care costs of diabetic patients in Spain. Diabetes Care 2004; 27:2616-21

² K. Nakamura et.al. Medical expenditure for diabetic patients: A 10-year follow-up study of National Health Insurance in Shiga, Japan. Public Health 2008; 122: 1226-1228

³ Jonsson B, CODE-2 Advisory Board. Revealing the cost of type II diabetes in Europe. Diabetologia 2002; 45: S5-12

⁴ 国立社会保障・人口問題研究所 平成 28 年度社会保障費用統計

⁵ JM. Chan et.al. Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men. Diabetes care 2002; 17: 961-9

⁶ Sone H, et.al. Energy intake and obesity in Japanese patients with type 2 diabetes. Lancet 2004; 17: 363(9404): 248-9

2・2 方法

1) アクサ健康保険組合・アクサ生命より提供を受ける健診データ・レセプトデータ、人事データを突合し、上記目的に沿った分析およびモデル構築を行う。分析およびモデル構築の詳細に関しては、以下の 2) を参照のこと。

2) 本研究の流れは以下の通りを予定する。

① 「個人情報の保護に関する法律」および「匿名データの作成・提供に係るガイドライン」等に基づき、個人情報を匿名化した上でデータの提供を受ける。匿名化は、アクサ健康保険組合・アクサ生命が作業を行う。

② データは厳重に管理した上で分析の際に使用する。

③ 健診、レセプトデータ及び人事データを突合し、ICD10コードの割り振られている代表的な生活習慣に関連した疾

患（例えば、糖尿病、高血圧症、高脂血症、動脈閉塞、脳血管障害、虚血性心疾患など）に対する医療費の集計分析を行う。この集計分析を通して、1元的、または2元的に医療費の主要因のあたりをつけ、後のモデリングのプロセスに活かす。

また、突合の際に、各レコードがうまく紐づかず、欠損値がいくつかのカラムで多く発生すると考えられる。そのような欠損値の取り扱いを統計的に処理が可能かどうかを検討する。特に、欠損の多いレコードに偏った分布がみられるかどうかを予め把握しておくことで、後のプロセスで得られる結果がデータ全体にイえることなのか、それとも特定の分布のみに結論づけることができるものなのかどうかを言えるようにする。

医療費に関しては、データ上では各傷病別に医療費が分配された状態になっていないため、例えば、糖尿病の医療費を適切に推計することができるかどうかを検討する。仮に適切に対象医療費の推計ができる場合は、推計医療費の平均値を、⑤における、介入による医療費の低減効果を見積もる際に利用する。

④ 医療費の要因分析を、機械学習の手法および従来の分析手法（一般化線形モデルなど）を利用し推定する。

特に、既存の分析ではあまり考慮されてこなかった交互作用に着目し、より細分化されたリスク要因を洗い出すことを目的とする。（一次報告）

⑤ 糖尿病の将来疾患リスクを算出するモデルを検討する。どのようなモデルが最適かは授受するデータに依存するが、混合効果モデルによる時系列モデリングや、多重代入法をもとにした潜在成長モデル、交互作用を考慮することができる機械学習モデル等を想定している。また、モデルの評価はAUCなどで行うことを想定している。予測対象となる目的変数は健診数値とし、例えば、糖尿病の場合は、主に空腹時血糖が126mg/dL以上である場合を糖尿病の疾患リスクありと判断する¹。分析にはRを使用し、データのサイズが膨大でRで管理しきれない場合は、RDBMSの採用を考える。

⑥ 将来疾患リスクをもつ被保険者に早期介入を行うことで、どれだけ医療費の削減効果があるかをシミュレーションし、介入効果を定量的に見積もる。（二次報告）

¹ 糖尿病診療ガイドライン 2016

実施・収集期間 1年間

研究期間

1年間

2・3 対象及び資料等

1) 対象

- ・アクサ健康保険組合の被保険者集団

2) 資料・情報

- ・アクサ健康保険組合の被保険者の健診データ
- ・アクサ健康保険組合の被保険者のレセプトデータ
- ・アクサ生命の人事データ

これらのデータは、個人情報保護法に照らして、アクサ健康保険組合から東京大学への委託（受託研究契約）によって提供され、東京大学で分析に使用する（添付資料受託研究申込書および受託研究契約書）。なお、本研究は健康保険組合が健康保険法の主旨に基づいて実施する業務に伴って実施されるものである。

<収集するデータ項目>

健診データ：

性別、年齢、BMI、既往歴の有無、収縮期血圧、中性脂肪、HDL、LDL、GOT、 γ -GTP等（約5年分）

レセプトデータ：

医療費合計、各種傷病の有無、年度等（約5年分）

人事データ：

所属部署、勤怠、通勤情報等（約7年分）

2・4 研究参加者の実体験

・特になし

3. 研究を実施する施設とその役割

1) 該当する施設名とその役割 (別途添付可)

- ① インフォームド・コンセントを受ける施設：アクサ健康保険組合、アクサ生命
 - ② 資料等を収集する施設：アクサ健康保険組合、アクサ生命
 - ③ 資料等を匿名化する施設：匿名化作業は、アクサ健康保険組合およびアクサ生命人事部が個別に作業
 - ④ 資料等を解析する施設：東京大学政策ビジョン研究センターデータヘルス研究ユニット伊藤国際学術センター4F
 - ⑤ 資料等を保存する施設：東京大学政策ビジョン研究センターデータヘルス研究ユニット伊藤国際学術センター4F
- 2) (該当する場合) 学外施設での対応とその状況
なし

4. 研究における倫理的配慮

4・1 インフォームド・コンセント

1) 実施方法

a. アクサ健康保険組合の場合はHPで以下のように公表済。

「匿名加工情報の作成と提供については以下のように実施します。

当健康保険組合では、保健事業や疫学調査等のために、匿名加工情報を継続的に作成し、電子的な通信手段もしくはDVD等の物理媒体を用いてレセプト分析業者・研究機関等に提供いたします。作成及び提供する匿名加工情報に含まれる情報の項目は、性別、生年月、医療保険の資格情報(加入時期、脱退時期、本人・家族区分等)、診療報酬明細書の受診履歴、健診の受診履歴です。なお、個人を特定できる情報は含まれておりません。」

b. アクサ生命の場合は社員に開示している「社員情報の取り扱いに関するポリシー」の第10条共同利用にて、社員に周知済み

2) 特に倫理的な配慮を必要とする研究参加者への配慮の有無と対応策 → ■なし

4・2 個人情報保護

1) 本学における個人情報の有無とその種類 → ■あり

匿名化された健診・レセプトデータ

2) 個人情報保護の方法

機微情報の匿名化处理

LANに接続していない状態のPCで厳重に保管

3) 研究期間終了後：個人情報の保存/廃棄方法

研究期間終了後、データは廃棄し記録メディアは返還する。

4・3 資料等の取扱

アクサ健康保険組合・アクサ生命より電子データの提供を受ける(紙媒体の資料はなし)。当研究ユニット研究室内の鍵のかかるキャビネット(引き出し等)内で、パスワードで制限され、LANに接続していない状態のディスクで厳重に保管する。提供を受けたデータは研究終了の2年後に廃棄する。

5. 安全の確保

1) 研究によって研究参加者に生じうる危険や不快等

匿名化されたデータを用いるので、研究によってデータを取得した本人(研究参加者)が危険や不快等を感じるおそれはない。

2) 危険や不快等への対応策

研究に伴って直接危険や不快を生じさせるおそれはないが、研究の結果、生活習慣病の疾患リスクの将来予測ができた場合、その実用化においてはリスクの高い者に対して配慮が必要となる可能性はある。ただし、本研究の直接の目的はモデルを構築することであり、被保険者個人に対してリスク予測の結果をフィードバックすることは考えていない。